E. Approaches to Process Control

3.1 Basic Control Concepts and Definitions

Central to the control of any entity or system is the concept of 'feedback'. On a formal basis, it is defined as the process whereby the difference between the output of a system and a reference input is used to maintain a prescribed relation between that output and the reference. Feedback is, of course, an everyday phenomenon. Relative to engineering, its use for the control of mechanical systems is generally considered to date from James Watt's invention of the ball-governor to control his newly developed steam engine.

Control systems can be broadly categorized according to the use that is made of feedback. Figure 3.1-1 shows a generally accepted scheme in which systems are classified as manual, closed-loop, or open-loop. Using manual control, a human being compares the output of the process to that desired and then adjusts the actuator so as to maintain that output as specified. The operator is using feedback in that he or she first monitors the instrumentation and then combines that information with knowledge of the plant dynamics to determine the appropriate signal for the actuator.

Closed-loop control is similar to manual operation except that the plant's output is used directly in the determination of the control signal. This is shown in the middle portion of the figure. The 'reference output' reflects the desired or specified manner in which the plant is to be operated. This is compared to the measured output which is the signal obtained from the instrumentation. The difference between these two quantities is the 'actuating error signal'. The control signal, which is sent to the actuator, is a function of this error. The error and control signals could be computed by either analog or digital computers. Or, as was the case with the ball-governor for the steam engine, they could be determined by mechanical means. Under closed-loop control, the signal to the actuator is generated as the process evolves. Hence allowance is automatically made for any perturbations that may affect the plant's dynamics. One final point that is worth noting about closed-loop control is that the control signal is a function of the measured rather than the actual output of the plant. Hence, if the instruments fail, so will the control sys-

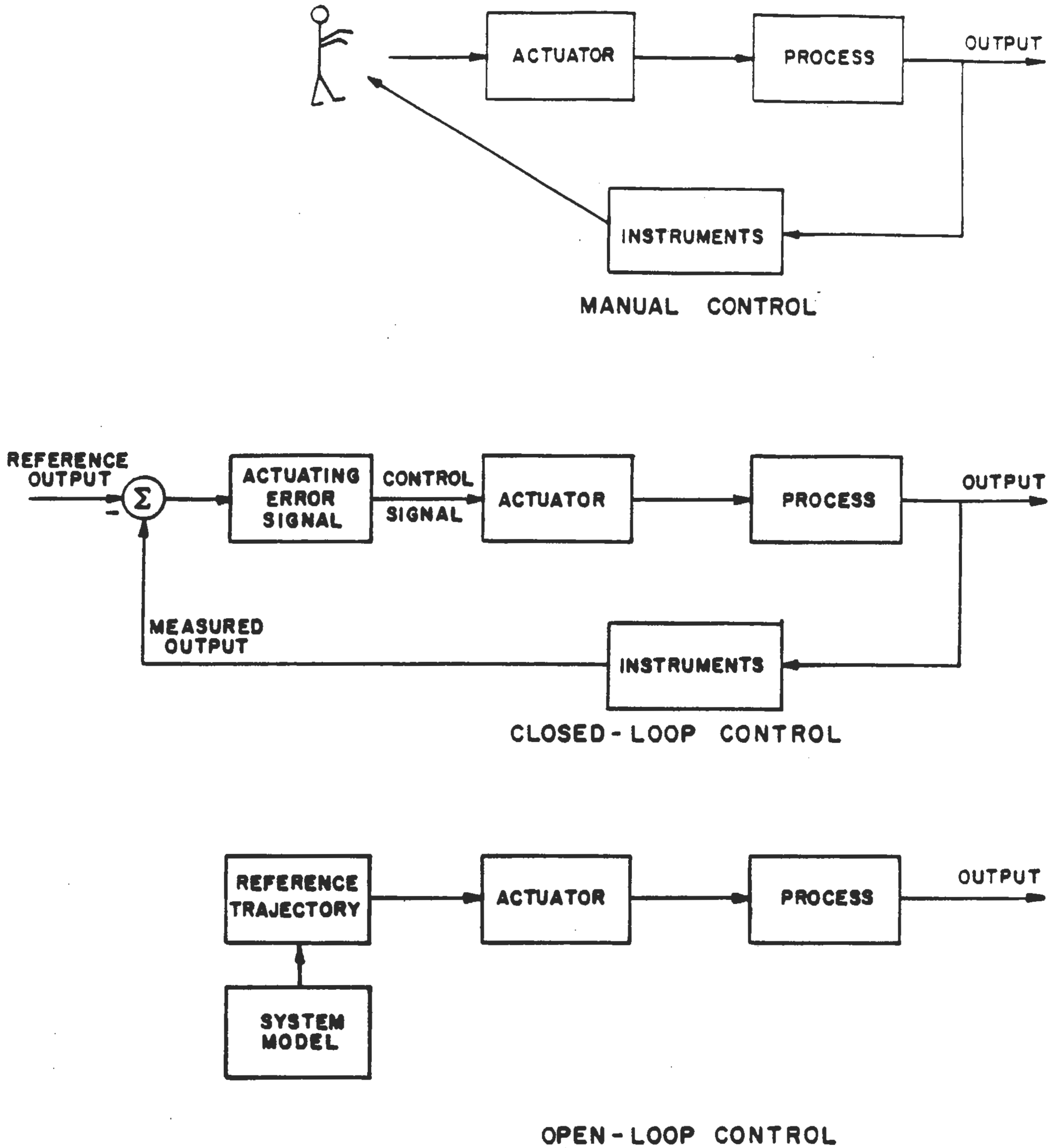

Figure 3.1-1 Manual, Closed-Loop, and Open-Loop Control

tem. This may not matter during manual operation because human beings can sometimes recognize and compensate for failed sensors. However, under closed-loop conditions, the sensor's erroneous output will be used. The solution to this problem is sensor validation which is discussed in sections 10.2.1 and 10.4.1 of this report.

Open-loop control is shown in the lower portion of the figure. The distinguishing feature of this approach is that no use is made of feedback. Rather, a predetermined reference trajectory is specified. For example, an accurate model of the process in question could be constructed and simulations performed until the sequence of control signals needed to generate the desired output was identified. Open-loop control may or may not result in the desired plant behavior. Clearly if the model used to determine the trajectory is in error, then so will the control action. However, even with a perfect model, the results might be poor. There is, for example, no way to allow for unanticipated changes in the plant. Despite its drawbacks, open-loop control is extensively studied and, as discussed below, some sophisticated control strategies use it.

3.2 Control Methodologies

Enumerated here are the major control methodologies that are now in use. As general references, the texts 'State Functions and Linear Control Systems' by Schultz and Melsa (McGraw-Hill, 1967), 'Control System Design: An Introduction to State-Space Methods' by Friedland (McGraw-Hill, 1986), 'Modern Control Engineering' by Ogata (Prentice-Hall, 1970) and 'Process Dynamics and Control' by Seborg, Edgar, and Mellichamp (John Wiley & Sons, 1989) are recommended. The definitions given in section 3.2.3.1 of this report are from the first of these.

3.2.1 Proportional-Integral-Derivative Control

Proportional-Integral-Derivative or P-I-D control is the most commonly used method for the automated operation of process systems. Using this approach, the system itself is treated as a 'black box'. That is, the control signal is determined solely by comparison of the measured and reference outputs. Information on the dynamics of the process, even if available, is not used. Figure 3.2.1-1 is a block diagram of the approach. The error signal is taken as the difference between the reference and measured outputs of the process. If the controller output is only proportional to this error, then the control action is termed 'proportional'. If the controller output also contains a term that is a function of the accumulated error, then it is designated as 'proportional-integral'. Finally, if the controller output consists of terms representing both the integral and the rate of change of the error signal as well as the error signal itself, then it is referred to as 'proportional-integral-derivative' or P-I-D.

It is instructive to examine the mathematics of the P-I-D control approach. For simplicity assume that the actual and measured plant outputs are identical. That is, the instruments are functioning perfectly. Suppose that the controller is strictly proportional and that

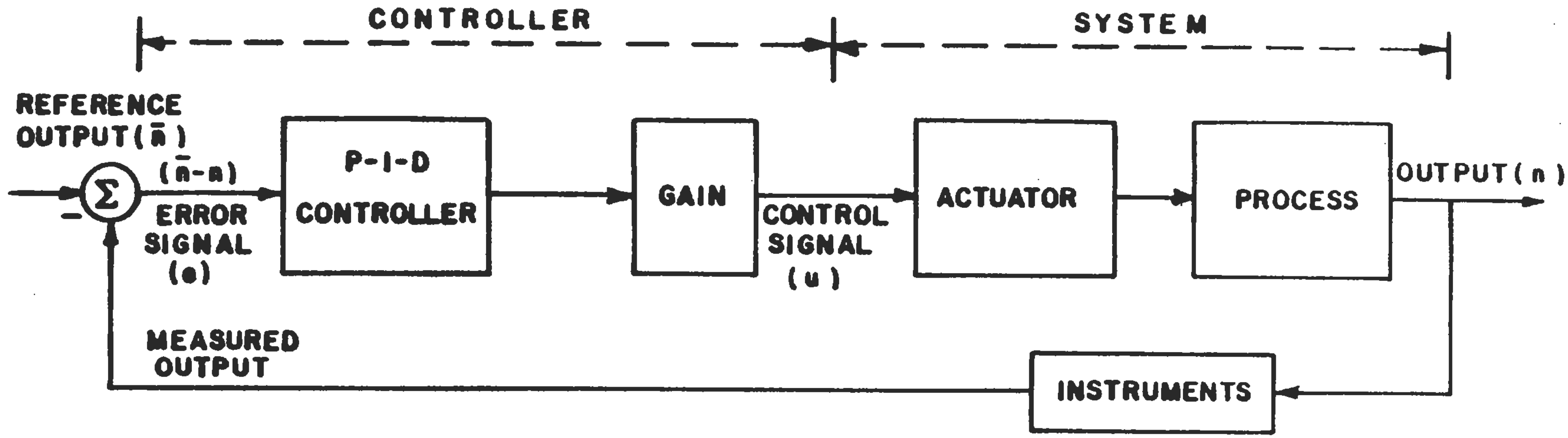

Figure 3.2.1-1 Proportional-Integral-Derivative Control

the system and the controller can be represented respectively by the following relations:

$$n(t) = Hu(t) \tag{3.2.1-1}$$

$$u(t) = Ke(t) \tag{3.2.1-2}$$

where: $n(t)$ is the plant output,
H denotes the process dynamics,
$u(t)$ is the control signal,
K is the gain of the controller, and
$e(t)$ is the error signal.

Also note that the error signal is defined as:

$$e(t) = (\bar{n}(t)-n(t)) \tag{3.2.1-3}$$

where $\bar{n}(t)$ is the reference output. Combining the above relations yields:

$$\begin{aligned} n(t) &= Hu(t) \\ &= HKe(t) \\ &= HK(\bar{n}(t)-n(t)) \end{aligned} \tag{3.2.1-4}$$

Solving for $n(t)$ yields:

$$n(t) = [\frac{HK}{1 + HK}]\bar{n}(t) \tag{3.2.1-5}$$

This simple relation illustrates the enormous advantages to using feedback and closed-loop control. Specifically, note that by making the gain arbitrarily large, the quantity $(HK/(1+HK))$ will be driven to unity and the observed plant output, $n(t)$, will become equal to the reference or desired output, $\bar{n}(t)$. Of special significance is that this result is achieved without any knowledge of the process dynamics and that it is valid for any reference output. Moreover, even if the process dynamics change or are perturbed, it should be possible to maintain the output of the process at the desired value. Hence the appeal of closed-loop control.

Unfortunately, it is not possible to achieve all of the theoretical advantages associated with the proportional controller described above. For one thing, the controller gain can not be made arbitrarily large. So doing will make the system increasingly sensitive to small perturbations and may eventually cause instability. This creates an immediate problem because, for finite values of K, the quantity $(HK/(1+HK))$ will be less than unity and hence $n(t)$ will never quite equal $\bar{n}(t)$. This is a major disadvantage of proportional control. Such systems exhibit an offset between the desired and observed outputs. This problem can be rectified by adding integral action to the controller. That is, instead of a control signal that is directly proportional to the error signal as in Equation 3.2.1-2, the control signal is of the form:

$$u(t) = Ke(t) + \frac{K}{T_i} \int_0^t e(t)dt \qquad (3.2.1\text{-}6)$$

where the quantity T_i is the integral time. The inverse of the integral time, which is termed the 'reset rate' has a physical interpretation. It is the number of occasions per minute that the proportional part of the control action is duplicated. The advantage of adding an integral term to the control action is that the control signal can have a non-zero value even though the error is zero. Thus, the use of integral action resolves the offset problem. Any difference between the desired and observed outputs will eventually cause the total error to accumulate to the point where it drives system response to the specified value. The drawback to integral action is that it may induce oscillations as the system converges on the specified output.

Controller performance may be further improved by adding a derivative term to the control signal. Thus,

$$u(t) = Ke(t) + \frac{K}{T_i} \int_0^t e(t)dt + KT_d \frac{de(t)}{dt} \qquad (3.2.1\text{-}7)$$

The quantity T_d is the derivative time. Physically, it is the interval by which the proportional part of the control action is advanced. Derivative action is anticipatory and thus its use may reduce oscillations. However, it may also amplify signal noise.

In summary, the advantages of the P-I-D approach are that the technique can be applied without regard to the process dynamics, that its basic concepts are readily understood, and that it can be implemented using either analog or digital equipment. Its disadvantages are that the selection of the control parameters (the gain and the integral and derivative times) is empirical and that these parameters will be valid for only one particular transient. Thus the use of the P-I-D approach for non-linear or time-delayed systems is impractical. Another drawback is that the technique can only handle one input and one output. Also, the theory of P-I-D control does not provide the basis for determining system stability.

3.2.2 Transfer Function Approach

The use of transfer functions in control system design represents a major advance over the P-I-D approach because the system is no longer treated as a black box. As a result, the stability of the controlled system can be analyzed and theoretically valid methods can be used to select the gain coefficient. However, whereas the P-I-D approach was in theory applicable to any process, use of the transfer function approach is limited to linear, time-invariant systems.

A transfer function is defined as the ratio of the Laplace transform of the system output to the Laplace transform of the system input. For a given system, it is obtained by first identifying the plant's describing differential equation, then taking the Laplace transform, and finally rearranging terms so as to obtain the ratio of the output to the input. This process results in a transfer from the time domain (differential equation) to the frequency domain (transfer function). As an illustration of the utility of the approach, suppose that the system (actuator plus process) shown in Figure 3.2.1-1 was actually described by the following second-order, linear differential equation:

$$u(t) = \ddot{y}(t) + \dot{y}(t) \qquad (3.2.2\text{-}1)$$

Taking the Laplace transform of this equation subject to the assumption that all initial conditions are zero yields:

$$U(s) = s^2Y(s) + sY(s) \qquad (3.2.2\text{-}2)$$

Hence, the transfer function of the process is:

$$\frac{Y(s)}{U(s)} = \frac{1}{s(s+1)} \qquad (3.2.2\text{-}3)$$

Further suppose that a control scheme is to be designed which uses the same feedback approach as did the P-I-D controller. That is, only the output of the system is fed back. The process is illustrated in Figure 3.2.2-1. The closed-loop behavior of the system is described by the relation:

$$[R(s)-Y(s)](K)[\frac{1}{s(s+1)}] = Y(s) \qquad (3.2.2\text{-}4)$$

Hence, the closed-loop transfer function is:

$$\frac{Y(s)}{R(s)} = \frac{K}{s^2+s+K} \qquad (3.2.2\text{-}5)$$

This can be written in factored form as:

$$\frac{Y(s)}{R(s)} = \frac{K}{(s-a)(s-b)} \qquad (3.2.2\text{-}6)$$

where, using the quadratic formula, the quantities a and b are determined to be:

$$a = -0.5+0.5\sqrt{1-4K}; \; b = -0.5-0.5\sqrt{1-4K} \qquad (3.2.2\text{-}7)$$

Finally, by taking the inverse Laplace transfer of (3.2.2-6), the closed-loop behavior of the system can be deduced. So doing,

$$y(t) = (\frac{1}{a-b})[ae^{at}-be^{bt}] \qquad (3.2.2\text{-}8)$$

Much can be determined about system stability and response using this

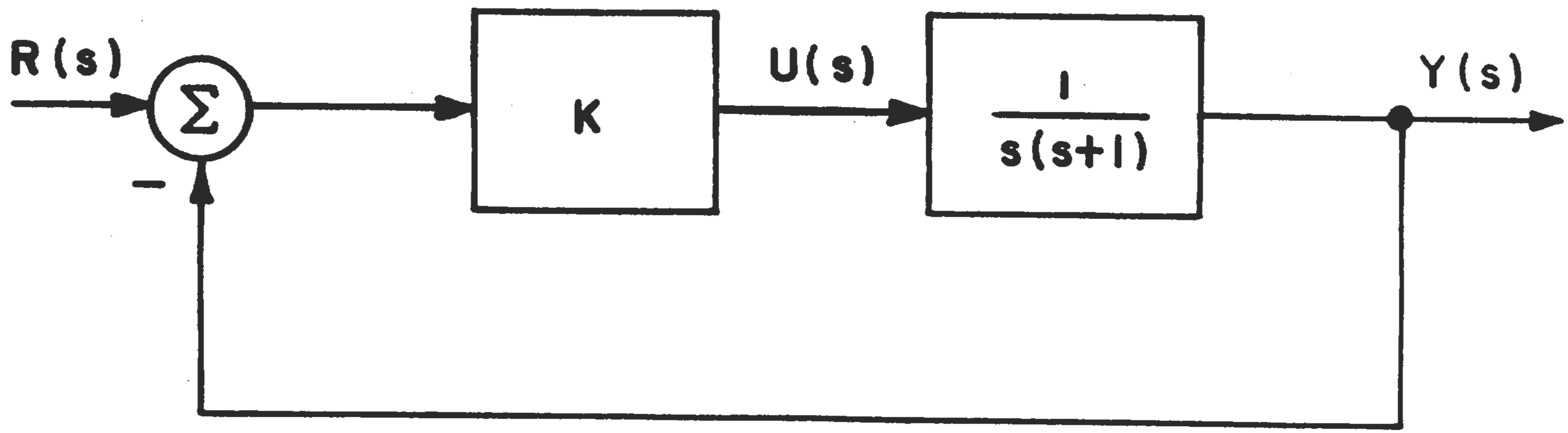

Figure 3.2.2-1 Transfer Function Approach to Controller Design

analytic solution for the behavior of the closed-loop system. First note that the value of the gain can range from zero to infinity. If the gain lies between 0.00 and 0.25, then the quantities a and b, which are the roots of the closed-loop transfer function, are negative real numbers. If the value of the gain exceeds 0.25, then those roots are complex numbers with negative real parts. Hence, for any value of the gain, the exponential terms (EXP(at) and EXP(bt)) will die out as time increases. The system is therefore unconditionally stable. (Note: For other systems, it is possible that for a certain range of gains, the roots would be positive implying unbounded exponential growth and instability.) Another important feature to note is that if the value of the gain exceeds 0.25, the system response will be under-damped meaning that the output will oscillate about its desired value before settling out. If the gain is equal to 0.25, the system should attain the desired response without overshoot. If the gain is less than 0.25, then there will still be no overshoot but the response time will be longer than is perhaps necessary.

The above example shows that by including a model of the system in the controller, much insight can be achieved. Using the P-I-D approach, the only guidance received was to make the gain as large as possible. But with the transfer function approach, it is possible to identify the range of gains for which the system will be stable and to select a gain within that range to achieve desired response characteristics. (Note: The system shown in Figure 3.2.2-1 was analyzed by obtaining the closed-loop response, Equation 3.2.2-8. In reality, this was not necessary because using methods such as 'root-locus' or Routh's stability criteria, the same information could have been obtained directly from the transfer function. This is a further strength of the technique.)

In summary, the transfer function approach can be applied to linear, time-invariant systems. It provides a rational means for analyzing system stability and for specifying certain system response characteristics. It can be implemented using either digital or analog equipment. Its disadvantages are that it is limited to linear systems, that it treats only single-input/single-output systems, and that it does not address certain fundamental issues such as what constitutes a proper choice of control signal. Also, the proper functioning of the controller is dependent on the accuracy of the model chosen as representing the system during the design stage. This raises a number of issues including model validation and maintenance.

3.2.3 State-Space Methods

State-space methods are arguably the most successful of the various control techniques now available. This methodology was the one used to design the control systems for the Apollo spacecraft and it is the one now used for many aerospace applications. The basic concept originated in the 1960s and, although nearly three decades have passed, the methodology is still referred to by the misnomer 'modern control theory'. As with the transfer function approach, the strength of state-space methods is the use of a system model. Only instead of using a single differential equation to relate system output to system input, a set of first order differential equations is used. Thus, full advantage is taken of the knowledge given by the internal structure of the system. For example, in the preceeding section of this report, a system characterized by the following equation was analyzed:

$$u(t) = \ddot{y}(t) + \dot{y}(t) \qquad (3.2.3\text{-}1)$$

This is a second order differential equation where y is the output and u is the input. Using the state-space approach, this is rewritten as two first order differential equations. The state variables, here denoted by the symbols x_1 and x_2, are defined as:

$$x_1 = y \qquad (3.2.3\text{-}2(a))$$

$$x_2 = \dot{y} \qquad (3.2.3\text{-}2(b))$$

Thus,

$$\dot{x}_1 = x_2 \qquad (3.2.3\text{-}3(a))$$

$$\dot{x}_2 = -x_2 + u \qquad (3.2.3\text{-}3(b))$$

The second of these equations is obtained by solving (3.2.3-1) for the quantity $\ddot{y}$ and then substituting x_2 for $\dot{y}$. Equations 3.2.3-3 can be written in matrix form to facilitate solution on a digital machine:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \qquad (3.2.3\text{-}4)$$

Note that by substituting (3.2.3-3(b)) into (3.3.2-3(a)) and then noting that $\dot{x}_1$ equals $\dot{y}$ and that $\dot{x}_2$ equals $\ddot{y}$, equation 3.2.3-1 is regenerated. Thus, both representations (3.2.3-1 and 3.2.3-3) provide descriptions of the system. The advantage associated with the state-space approach is that by using a set of n first order differential equations to describe an n<u>th</u> order system, the internal dynamics of that system become accessible. Figure 3.2.3-1 is a block diagram illustrating the value of this approach for controller design. Instead of feeding back merely the output of the process, each state variable is assigned a feedback coefficient. (<u>Note</u>: The symbol 'h' is used to denote the feedback coefficients in the figure.) Hence, for an n<u>th</u> order system, n degrees of freedom are introduced. This provides the control engineer with enormous flexibility. In particular, by the judicious choice of the feedback or gain coefficients associated with each of the state variables, the shape of the system's response as well as its stability can be specified. There are other advantages as well to the use of state-space methods. Using this approach, controllers can be designed for systems that have multiple inputs and out-

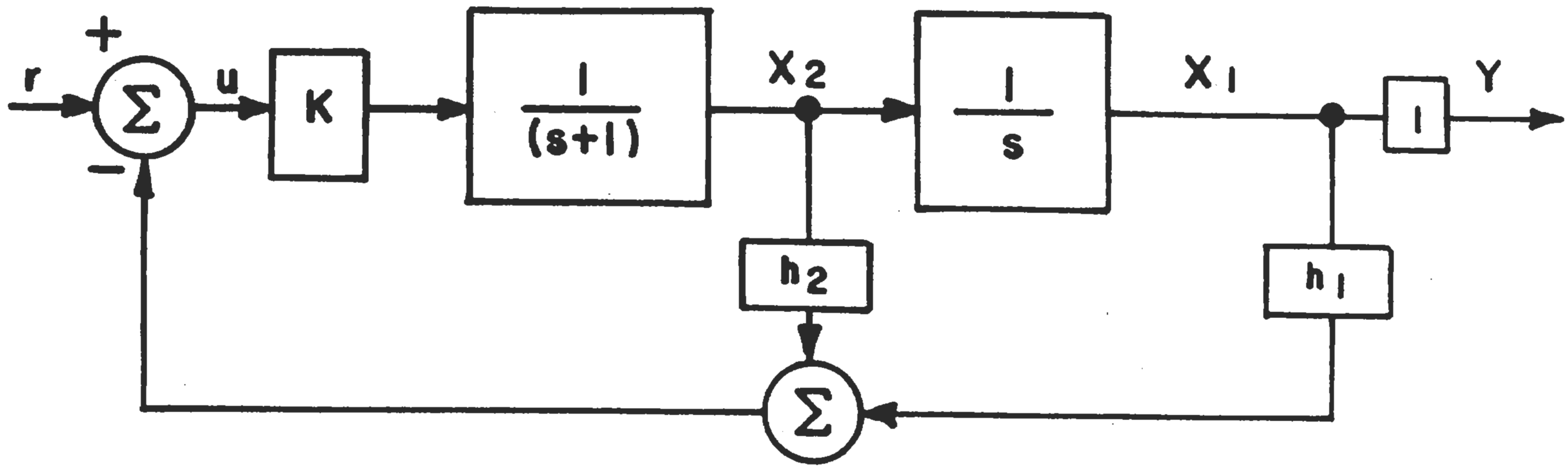

Figure 3.2.3-1 State-Space Approach to Controller Design

puts. Also, as is discussed below in section 3.2.3.1 of this report, the state-space methodology addresses fundamental issues such as the criteria by which it can be determined if a system is in fact controllable. The technique is however not without its drawbacks. For one thing, its use is limited to linear systems. For another, it may not be possible to give a physical meaning to each of the state variables and even if it is, there is no guarantee that the variable can be measured. For example, relative to a nuclear reactor, precursor concentrations are state variables that cannot be directly observed.

3.2.3.1 Controllability and Observability

The concepts of 'controllability' and 'observability' were first introduced by Kalman in the 1950s. The former concerns the capability of the control signal (i.e., the input to the controller) to affect each of the state variables. The latter is the capability of each state variable to affect the output. On a formal basis, these are defined as:

(1) Controllability: A system is said to be 'controllable' if any initial state $x(0)$ can be transferred to any final state $x(t_f)$ in a finite time, $t_f \geq 0$, by some control, u.

(2) Observability: A system is said to be 'observable' if every state $x(0)$ can be exactly determined from measurements of the output y over a finite interval of time, $0 \leq t \leq t_f$.

Most physical systems are both controllable and observable. Hence the need to verify these properties often appears to be of only academic interest. However, such should not be the case, especially when considering complex systems. Two issues are of concern. First, while it is true that most physical systems are controllable, their associated mathematical models may lack that property. This most often occurs when a linearized model is used to represent a non-linear system. The second issue is that some systems may be only partially controllable. This is a complex problem that has not yet been fully addressed on a theoretical basis. Specifically, the formal definition of controllability provides no information on the degree to which a system may or may not be controllable. This would be most desirable because many systems and their associated control mechanisms fulfill the definition of controllability in the sense that given enough time and with no restriction on the allowed trajectories, it is possible to cause the desired change using the specified control device. Yet, the resulting control action may be of poor quality. For example, suppose that a controller for a spacecraft nuclear reactor is to be constructed in which the reactor power is to be adjusted in response to a temperature sensor. Moreover, suppose that this sensor must be located some distance from the reactor so that radiation damage effects are minimized. Is the system controllable? Clearly, if the reactor power were to change, then so would the temperature of the reactor coolant. But the sensor is located on a boom and hence the temperature of the coolant in which it is immersed will not change until heat has been conducted along the entire length of the boom.

Thus, there will be a significant time delay associated with the use of such a sensor for control and the resulting control action may be oscillatory. The system is controllable according to the definition. Nevertheless, the resulting power trajectory may exhibit undesirable features.

Concerns related to controllability and in particular to the need to refine the concept so that the controllor is capable of producing not only the desired end-state but also an acceptable trajectory have been a major aspect of the MIT program in the closed-loop digital control of reactor power. Much has been accomplished. In particular, the concept of 'feasibility of control' has been defined and the rate of change of reactivity (as opposed to the reactivity itself) has been recognized as the appropriate means of control. Both of these concepts are discussed in detail in Chapter Four of this report. For the present, it is sufficient to note that 'feasibility of control' is, like controllability, a property of the controller. Observance implies that the allowed states of a system are restricted such that it will always be possible to halt a transient on demand. Relative to nuclear reactors, completed work at MIT has provided a practical method for both defining and incorporating such a property in controllers for reactor neutronic power. The selection of the rate of change of reactivity as the control signal bears on this issue as well. Specifically, it is a directly controllable quantity in the sense that it can be altered on demand. In contrast, reactivity can only be changed over some finite interval of time.

3.2.3.2 Summary of State-Space Approach

In summary, the state-space approach is currently the most widely used technique for the design of controllers for complex systems. Its advantages are that it can be used both to assure system stability and to achieve a desired system trajectory. Moreover it is capable of handling multiple-input/multiple-output systems. Also, its underlying theory addresses fundamental issues such as what constitutes a controllable system. Its disadvantages are that its use is restricted to linear systems, that its implementation may require information on variables that cannot be measured, and that it does not address the degree to which a system may be controllable. Also, the solution of most state-space design problems requires the use of a digital computer.

3.2.4 Optimal Control

Control system performance is considered optimal in the sense that the resulting trajectory either minimizes or maximizes a designer-specified performance index. For example, it may be desired to complete a transition in minimum time or with as little fuel consumption as possible. The realization of optimal control is achieved by using a state-space representation. That is, the system (process plus actuator) is described using a set of first order differential equations. In the state-space approach, each state variable (i.e., the independent variables in the describing equations) is assigned a feedback coefficient or gain. The user then selects the desired shape or trajectory of the system's response and solves for the gains so as to achieve that response. The burden of selecting the appropriate trajectory is therefore on the designer. With optimal control, the designer merely selects a performance criteria. The system's gain is chosen so as to minimize (or, if appropriate, maximize) this index. Thus, the designer is relieved of a major burden. Unfortunately, another one is imposed in that the mathematics of the approach are often intractable. For example, if the Pontryagin approach to optimal behavior is used, then it becomes necessary to solve a set of partial differential equations with split boundary conditions. That is, some conditions are known that apply to the initiation of the transient and some to its termination, but there is not a complete set at either the start or end of the transient. Hence, the problem must be solved by iteration. Figure 3.2.4-1 illustrates the problem. The objective is to identify a trajectory that optimizes some performance index. Assume that the optimal trajectory is the one indicated by the broken line. The curves denoted by the letters 'a' and 'b' represent successive attempts to identify the optimal solution. The separate pieces of each curve are generated by marching forward from the initial conditions and backwards from the final ones with the hope of the two pieces joining to form a single path. This process is iterative often with many successive attempts being required before the desired path is found. Thus, the solution to an optimal control problem is calculation-intensive. This means that, unless the problem is very simple, the calculations will take longer than the time that is available during each sampling interval. Hence, many optimal control problems are solved in their entirety prior to initiating the actual control action. The solution to an optimal control problem is therefore usually not a closed-form law but rather a sequence of signals that are to be applied at each sampling interval. The drawback to this approach is that this sequence will be valid only for the boundary conditions that were initially specified. No provisions exist for variations in those conditions and there is no continuous feedback of measured signals. In other words, this is a form of open-loop control. If the system model is perfect and if the boundary conditions have not changed, then the resulting control action will be as desired. Otherwise, an incorrect trajectory will be generated. Alternatives to the split boundary condition problem as formulated by Pontryagin do exist. For example, Bellman's dynamic programming avoids that problem. However, it too is calculation-intensive. This is the weakness of the optimal approach to system control.

A major accomplishment of the MIT-SNL program that is described in this report is that a method was found to obtain closed-loop, time-optimal control laws for the adjustment of a reactor's neutronic power. Details are given in Chapters Four and Twelve of this report.

In summary, the advantage of optimal control is that it allows the control engineer to obtain a desired response without having to specify the details of the trajectory. (Note: In the language of the state-space approach, "without having to select the locations of the closed-loop poles.") Another advantage is that the technique is

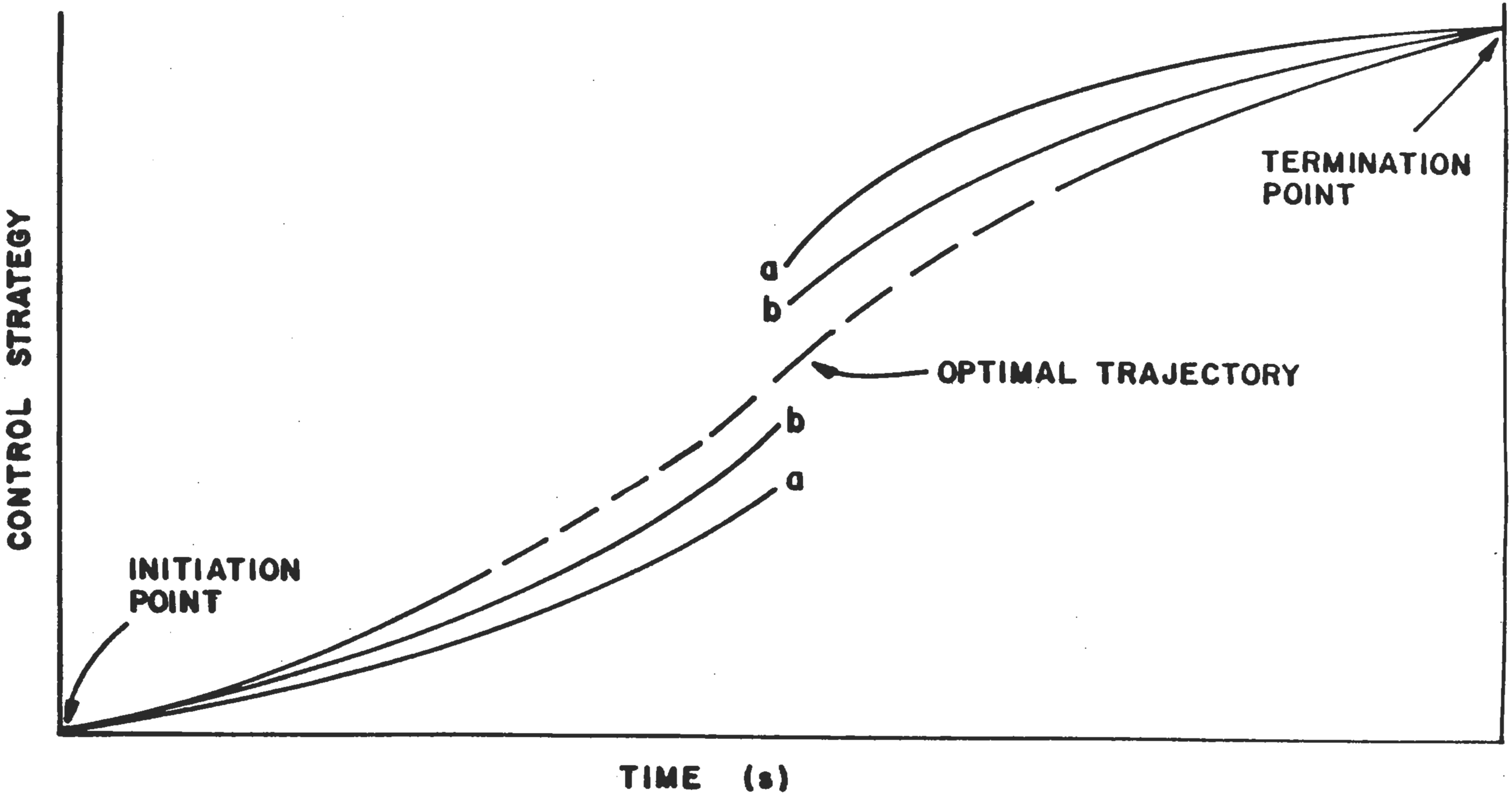

Figure 3.2.4-1 Successive Approximations to an Optimal Path

applicable to non-linear systems. Disadvantages are that the mathematics associated with the technique are often intractable and this in turn may cause resulting control action to be open-loop with no provisions for feedback.

3.3 Trends and Unresolved Issues in Control System Design

The brief review of control methodologies given in section 3.2 of this report considered the P-I-D approach, the use of transfer functions, state-space methods, and optimal control. There are of course many variations on these basic techniques as well as other distinct methodologies. However, the ones examined here are the dominant approaches. Several trends and unresolved issues are apparent from the review. These bear directly on the design of control systems for complex technologies such as nuclear reactors.

First, increasing reliance is being placed upon system models. The P-I-D approach avoids use of a model. Transfer functions are themselves a model with only the system input and output accessible. Both state-space and optimal control assume the existence of accurate models replete with details of the system's inner structure. As a general observation, it appears that the mathematics available for utilizing a model has outstripped the capability to delineate such models. That is, the limiting factor in taking advantage of the benefits of advanced control concepts is, in many respects, the difficulty associated with obtaining accurate representations of the plant. In particular, if these control techniques are to be applied to systems for which safety is of paramount importance, then issues such as model validation, calibration, and maintenance should be addressed on a formal basis.

A second issue is that of partial controllability. This was discussed in section 3.2.3.1 of this report. The issue is of particular importance in the control of systems that are either non-linear or time-delayed. In such systems the influence of one variable upon another may vary substantially as the result of small changes in the control signal.

A third issue is that of non-linear control. Most control methodologies are intended for use on linear systems. Moreover, even if a particular control approach is capable of treating the non-linear case, it is often applied to a linearized model in order to minimize mathematical complexity. This practice is certainly acceptable for static situations in which the objective of the control action is to maintain a specified condition. But it is not acceptable for transients. The now decade-old revolution in digital technology has made it possible to apply enormous computing power to individual control loops. Advantage should be taken of this new resource to design controllers in terms of the non-linear system.